IMM 358
April 1967

Courant Institute of Mathematical Sciences

On the Estimation of Location Parameters in the Multivariate One-Sample and Two-Sample Problems

Madan Lal Puri and Pranab Kumar Sen

Prepared under a Sloan Foundation Grant
for Statistics, and Research Grant
GM-12868 with the National Institutes
of Health, Public Health Service

New York University

IMM 358
April 1967

New York University
Courant Institute of Mathematical Sciences

ON THE ESTIMATION OF LOCATION PARAMETERS IN THE MULTIVARIATE ONE-SAMPLE AND TWO-SAMPLE PROBLEMS

Madan Lal Puri[1] and Pranab Kumar Sen[2]

[1] This report represents results obtained in part while this author was at the University of North Carolina, Chapel Hill.

[2] Presently at the University of North Carolina, Chapel Hill. On leave of absence from the University of Calcutta.

This work was partially supported by a Sloan Foundation Grant for Statistics, and the National Institutes of Health, Public Health Service, Grant GM-12868.

1. Introduction and Summary

Estimates of location parameters based on rank order tests in the univariate one-sample and two-sample problems have been considered by various authors. For the details of references, the reader is referred to the most recent papers of Gastwirth [4] and Bloch and Gastwirth [3]. In this paper we consider the estimates of location parameters in the multivariate one-sample and two-sample problems. These estimates are obtained through the use of coordinate wise rank order test statistics such as the Wilcoxon or normal scores statistic considered by the authors in [9] and [11]. The distribution of these estimates is shown to be symmetric with respect to the parameters being estimated. These estimates are also translation-invariant and robust. Furthermore the estimates are shown to be asymptotically normal. Finally the asymptotic efficiencies of these estimates relative to the estimates based on the vectors of means and medians are obtained by applying the criterion of Wilks generalized variance (cf. Anderson [1], p. 166).

2. Point Estimates

Let $\underset{\sim}{X}_\alpha = (X_{1\alpha},\ldots,X_{p\alpha})$, $\alpha = 1,\ldots,m$ and $\underset{\sim}{Y}_\beta = (Y_{1\beta},\ldots,Y_{p\beta})$, $\beta = 1,\ldots,n$ be independent vector values random variables drawn from p-variate absolutely continuous cdfs (cumulative distribution functions $F(\underset{\sim}{x})$ and $G(\underset{\sim}{x}) = F(\underset{\sim}{x} - \Delta)$ respectively where $\underset{\sim}{x} = (x_1,\ldots,x_p)$ and $\underset{\sim}{\Delta} = (\Delta_1,\ldots,\Delta_p)$. For each coordinate consider the statistic

$$(2.1) \qquad h_j = h_j(X_{j1},\ldots,X_{jm};\ Y_{j1},\ldots,Y_{jn}) = \sum_{\alpha=1}^{N} E_{N\alpha}^{(j)}\ Z_{N\alpha}^{(j)}/m$$

where $E_{N\alpha}^{(j)}$ is the expected value of the α-th order statistic of a sample of size $N = m+n$ from a distribution Ψ_j; $Z_{N\alpha}^{(j)} = 1$ if the α-th smallest observation from the combined sample $(X_{j1},\ldots,X_{jm};\ Y_{j1},\ldots,Y_{jn})$ is an X observation and $Z_{N\alpha}^{(j)} = 0$ otherwise, $\alpha = 1,\ldots,N$; $j = 1,\ldots,p$. Let for each $j = 1,\ldots,p$

$$(2.2) \qquad \Delta_j^* = \sup\left\{\Delta_j : h_j(X_{j1},\ldots,X_{jm};\ Y_{j1}-\Delta_j,\ldots,Y_{jn}-\Delta_j) > \mu_j\right\}$$

$$(2.3) \qquad \Delta_j^{**} = \inf\left\{\Delta_j : h_j(X_{j1},\ldots,X_{jm};\ Y_{j1}-\Delta_j,\ldots,Y_{jn}-\Delta_j) < \mu_j\right\}$$

where μ_j is a point of symmetry of the distribution of h_j when $\Delta_j = 0$, $j = 1,\ldots,p$. It was shown in [6] that the estimate $\hat{\Delta}_{N,j} = (\Delta_j^* + \Delta_j^{**})/2$ of Δ_j has more robust efficiency than the classical estimate $\sum_{\beta=1}^{n} Y_{j\beta}/n - \sum_{\alpha=1}^{m} X_{j\alpha}/m$ in the corresponding univariate problem. Thus, as a natural extension of [6], we propose

$$\hat{\underset{\sim}{\Delta}}_N = (\hat{\Delta}_{N1}, \ldots, \hat{\Delta}_{Np}) \tag{2.4}$$

as an estimate of $\underset{\sim}{\Delta}$ for suitable functions

$$\underset{\sim}{h} = (h_1, \ldots, h_p) \; . \tag{2.5}$$

Now let us consider the one-sample problem. Suppose $\underset{\sim}{Z}_\alpha = (Z_{1\alpha}, \ldots, Z_{p\alpha})$, $\alpha = 1, \ldots, N$ is an independent sample from a p-variate cdf $F(\underset{\sim}{x} - \underset{\sim}{\theta}) = F(x_1 - \theta_1, \ldots, x_p - \theta_p)$ where F is absolutely continuous and is diagonally symmetric about zero. (Let us recall that F is diagonally symmetric about zero, if its density remains invariant under simultaneous changes of signs of all the coordinate variates (cf. [11]).) As before consider for every univariate sample $X^{(j)} = (Z_{j1}, \ldots, Z_{jN})$, a statistic

$$h_j^* = h_j^*(Z^{(j)}) = \sum_{\alpha=1}^{N} E_{N\alpha}^{(j)} \, Z_{N\alpha}^{(j)}/N \; , \qquad j = 1, \ldots, p \tag{2.6}$$

where $E_{N\alpha}^{(j)}$ is the expected value of the α-th order statistic of a sample of size N from a distribution $\Psi_j^*(x)$ given by

$$\Psi_j^*(x) = \Psi_j^{**}(x) - \Psi_j^{**}(-x) \quad \text{for} \quad x \geq 0 \quad \text{and} \quad \Psi_j^*(x) = 0, \text{ otherwise.} \tag{2.7}$$

$\Psi_j^{**}(x)$ is symmetric about zero or uniform over (-1,1,), and $Z_{N\alpha}^{(j)} = 1$ if $Z_{j\alpha} > 0$ and $Z_{N\alpha}^{(j)} = 0$ otherwise; $\alpha = 1, \ldots, N$; $j = 1, \ldots, p$. Denote

$$(2.7) \qquad \theta_j^* = \sup\{\theta_j : h_j^*(Z_{j1}-\theta_j,\ldots,Z_{jN}-\theta_j) > \mu_j\}$$

$$(2.8) \qquad \theta_j^{**} = \inf\{\theta_j : h_j^*(Z_{j1}-\theta_j,\ldots,Z_{jN}-\theta_j) < \mu_j\}$$

where μ_j is a point of symmetry of h_j^* when $\theta_j = 0$. Then for suitable functions h_j^*, we propose

$$(2.9) \quad \hat{\underset{\sim}{\theta}}_N = (\hat{\theta}_{N1},\ldots,\hat{\theta}_{Np}), \quad \hat{\theta}_{Nj} = (\theta_j^* + \theta_j^{**})/2 , \qquad j = 1,\ldots,p$$

as an estimate of $\underset{\sim}{\theta}$.

The estimates $\hat{\underset{\sim}{\Delta}}_N$ as well as $\hat{\underset{\sim}{\theta}}_N$ defined above form a general class of estimates. Two important members of this class (to which we shall pay specific attention in Section 5 when we discuss the relative asymptotic performances of the estimates) are (i) the Wilcoxon type estimates $\hat{\underset{\sim}{\Delta}}_{N(R)}$ and $\hat{\underset{\sim}{\theta}}_{N(R)}$ resulting from the Wilcoxon statistic by taking for Ψ_j and Ψ_j^* respectively the rectangular distribution over (0,1) and (ii) the normal scores type estimates $\hat{\underset{\sim}{\Delta}}_{N(\Phi)}$ and $\hat{\underset{\sim}{\theta}}_{N(\Phi)}$ resulting from the normal scores statistic by taking for Ψ_j or Ψ_j^{**} the standard normal distribution. The present authors investigated the properties of the tests based on the statistics (2.1) and (2.6) for the p-variate one-sample and c-sample problems in [11] and [9] respectively. This paper examines the properties of the estimates $\hat{\underset{\sim}{\Delta}}_N$ and $\hat{\underset{\sim}{\theta}}_N$ based on the statistics (2.1) and (2.6) respectively.

3. Regularity Properties

The following theorems are immediate extensions of the theorems proved by Hodges and Lehmann [6] for the case p = 1 and are therefore stated without proofs.

THEOREM 3.1. *The distribution of the estimate* $\hat{\underset{\sim}{\Delta}}_N$ *as well as that of* $\hat{\underset{\sim}{\theta}}_N$ *is (absolutely) continuous if* F *is (absolutely) continuous.*

THEOREM 3.2. *Invariance Properties.*

$$(3.1)\quad \hat{\underset{\sim}{\Delta}}_N(\underset{\sim}{X}_1,\ldots,\underset{\sim}{X}_m;\ \underset{\sim}{Y}_1+\underset{\sim}{a},\ldots,\underset{\sim}{Y}_n+\underset{\sim}{a}) = \hat{\underset{\sim}{\Delta}}_N(\underset{\sim}{X}_1,\ldots,\underset{\sim}{X}_m;\ \underset{\sim}{Y}_1,\ldots,\underset{\sim}{Y}_n)+\underset{\sim}{a}$$

$$(3.2)\quad \hat{\underset{\sim}{\theta}}_N(\underset{\sim}{Z}_1+\underset{\sim}{a},\ldots,\underset{\sim}{Z}_N+\underset{\sim}{a}) = \hat{\underset{\sim}{\theta}}_N(\underset{\sim}{Z}_1,\ldots,\underset{\sim}{Z}_N)+\underset{\sim}{a}$$

where $\underset{\sim}{a} = (a_1,\ldots,a_p)$ *is any* $1\times p$ *vector of constants.*

THEOREM 3.3. *The distribution of the estimate* $\hat{\Delta}_N$ *defined in* (2.4) *is diagonally symmetric about* $\underset{\sim}{\Delta}$ *if either of the following two conditions hold:*

(a) *The distribution* F *is diagonally symmetric, and for each* $j = 1,\ldots,p$

$$(3.3)\quad h_j(X_{j1},\ldots,X_{jm};\ Y_{j1},\ldots,Y_{jn})+h_j(-X_{j1},\ldots,-X_{jm};\ -Y_{j1},\ldots,-Y_{jn}) = 2\mu_j$$

and

$$(3.4)\quad h_j(X_{j1}+a,\ldots,X_{jm}+a;\ Y_{j1}+a,\ldots,Y_{jn}+a) = h_j(X_{j1},\ldots,X_{jm};\ Y_{j1},\ldots,Y_{jn}+a \quad \textit{for all } a.$$

(b) The sample sizes m and n are equal and for each $j = 1,\ldots,p$; h_j satisfies

$$(3.5)\quad h_j(X_{j1},\ldots,X_{jm};\ Y_{j1},\ldots,Y_{jn}) + h_j(Y_{j1},\ldots,Y_{jn};\ X_{j1},\ldots,X_{jm}) = 2\mu_j$$

and (3.4).

THEOREM 3.4. The distribution of $\hat{\underset{\sim}{\theta}}_N$ defined in (2.9) is diagonally symmetric about $\underset{\sim}{\theta}$ if

(i) F is symmetric about zero

(ii) $h_j^*(Z^{(j)}) + h_j^*(-Z^{(j)}) = 2\mu_j$; $j = 1,\ldots,p$

Remark: The conditions on h_j and h_j^* in Theorems 3.3 and 3.4 are satisfied both for the Wilconxon and Normal Scores statistics.

THEOREM 3.5. For every vector $\underset{\sim}{a} = (a_1,\ldots,a_p)$

$$(3.6)\quad P[h_j(X_{j1},\ldots,X_{jm};\ Y_{j1}-a_j,\ldots,Y_{jn}-a_j) < \mu_j;\ j = 1,\ldots,p]$$

$$\leq P[\hat{\underset{\sim}{\Delta}}_N < \underset{\sim}{a}] \leq P[h_j(X_{j1},\ldots,X_{jm};\ Y_{j1}-a_j,\ldots,Y_{jn}-a_j) \leq \mu_j;\ j = 1,\ldots,p]$$

and

$$(3.7)\quad P[h_j^*(Z_{j1}-a_j,\ldots,Z_{jN}-a_j) < \mu_j;\ j = 1,\ldots,p] \leq P[\hat{\underset{\sim}{\theta}}_N < \underset{\sim}{a}]$$

$$\leq P[h_j^*(Z_{j1}-a_j,\ldots,Z_{jN}-a_j) \leq u_j;\ j = 1,\ldots,p]$$

4. Asymptotic Normality

Let m(N) and n(N) for N = 1,2,... be a sequence of pairs of sample sizes tending to infinity in such a way that $m(N)/N \longrightarrow \lambda$, say, and let $\underset{\sim}{\Delta}_N$ be a sequence of the values of the parameter $\underset{\sim}{\Delta}$. Also for the one-sample problem consider the sequence of sample sizes N = 1,2,... and let $\underset{\sim}{\theta}_N$ be a sequence of values of $\underset{\sim}{\theta}$. In both the cases we shall indicate the dependence of h_j, h_j^* and μ_j by writing h_{Nj}, h_{Nj}^* and μ_{Nj} respectively. Let us denote the marginal distribution of $Y_{j\alpha}$ (or $Z_{j\alpha}$) by $F_j(x)$ and that of $(Y_{j\alpha}, Y_{k\alpha})$ (or $(Z_{j\alpha}, Z_{k\alpha})$) by $F_{j,k}(x,y)$ when $\underset{\sim}{\Delta} = \underset{\sim}{0}$ (or $\underset{\sim}{\theta} = 0$) and assume that the marginals are absolutely continuous. Next let us define

$$(4.1) \qquad d_{j,k} = \begin{cases} \int_0^1 J_j^2(x)dx - \left(\int_0^1 J_j(x)dx\right)^2 & \text{if } j = k; \\ \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} J_j[F_j(x)]J_k[F_k(y)]dF_{j,k}(x-y) - \int_0^1 J_j(x)dx \int_0^1 J_k(y)dy & \end{cases}$$

$$\text{if } j \neq k; \; J_j = \Psi_j^{-1}, \; j = 1,\dots,p \; .$$

$$(4.2) \qquad e_j = \int_{-\infty}^{\infty} \{dJ_j[F_j(x)]/dx\} \, dF_j(x) \; ; \quad J_j = \Psi_j^{-1}$$

$$(4.3) \qquad f_{j,k} = \begin{cases} \frac{1}{4} \int_0^1 J_j^{*2}(x)dx & \text{if } j = k \\ \frac{1}{4} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} J_j^*[2F_j(x)-1]J_k^*[2F_k(y)-1]dF_{j,k}(x,y) & \text{if } j \neq k \end{cases}$$

$$J_j^* = \Psi_j^{*-1} \; , \; j = 1,\dots,p \; .$$

$$(4.4) \qquad g_j = \int_0^{\infty} \{dJ_j^*[2F_j(x) - 1]/dx\} \, dF_j(x) \; ; \; J_j^* = \Psi_j^{*-1}$$

Finally, let $\underset{\sim}{a} = (a_1,\ldots,a_p)$ be a $1 \times p$ vector of constants and let $c_1,\ldots,c_N,\ldots$ be real constants and let

$$(4.5) \qquad \underset{\sim}{\Delta}_N = - \underset{\sim}{a}/c_N \quad \text{or} \quad \underset{\sim}{\theta}_N = - \underset{\sim}{a}/c_N$$

Then to prove the asymptotic normality of $\hat{\underset{\sim}{\Delta}}_N$ and $\hat{\underset{\sim}{\theta}}_N$ we shall make use of the following two theorems, the first of which is an immediate consequence of Theorem 6.1 of Puri and Sen (1966) and the proof of the second is given in [11].

THEOREM 4.1. *If* h_j, $j = 1,\ldots,p$ *are given by* (2.1) *with* Ψ_j, $j = 1,\ldots,p$ *satisfying the assumptions of Theorem* 6.1 *of Puri and Sen* (1966) *and if* $m(N)/N \to \lambda$ *as* $N \to \infty$, *then the random vector* $N^{1/2}(h_{Nj} - \mu_{Nj},\ j = 1,\ldots,p)$ *has asymptotically a p-variate normal distribution with mean vector zero and covariance matrix* $D = ((d_{jk}))$ *given by* (4.1). [*Here* $\mu_{Nj} = \int_{-\infty}^{\infty} J_j[F_j(x)]dF_j(x)$.]

THEOREM 4.2. *If* h_j^*, $j = 1,\ldots,p$ *are given by* (2.6) *with* Ψ_j^* *satisfying the assumptions of Theorem* 4.1 *of Sen and Puri* (1967), *then the random vector* $N^{1/2}(h_{Nj}^* - \mu_{Nj}^*,\ j = 1,\ldots,p)$ *has asymptotically a p-variate normal distribution with mean vector zero and covariance matrix* $F = ((f_{jk}))$ *given by* (4.3). [$\mu_{Nj} = \int_0^{\infty} J_j^*[2F_j(x) - 1]dF_j(x)$.]

Remark: (1) The limiting distribution of $N^{1/2}(h_{Nj} - \mu_{Nj}$, $j = 1,\dots,p)$ is nonsingular if and only if the functions J_j and F_j are such that the moment matrix of $J_i[F_j(x)]$, $j = 1,\dots,p$ is nonsingular. It is singular if and only if, a.s. F

$$(4.6) \qquad J_j[F_j(x)] = \sum_{k \neq j} \alpha_k F_k(x) + \text{constant [cf. [9]]}$$

(α_k are some constants)

(2) The limiting distribution of $N^{1/2}(h^*_{Nj} - \mu_{Nj}$, $j = 1,\dots,p)$ is nonsingular if and only if the moment matrix of $\{J^*_j[2F_j(x) - 1]$, $j = 1,\dots,p\}$ is nonsingular. It is singular, if and only if, a.s. F

$$(4.7) \qquad J^*_j(2F_j(x) - 1) = \sum_{k \neq j} \alpha_k J^*_k[2F_k(y) - 1] + \text{constant [cf. [11]}$$

(α_k are some constants) .

The following theorems give the relationship between the limiting distributions of $\underset{\sim}{h}_N$ and $\hat{\underset{\sim}{\Delta}}_N$; and $\underset{\sim}{h}^*_N$ and $\hat{\underset{\sim}{\theta}}_N$.

THEOREM 4.3. Let G be the continuous p-variate distribution function whose marginals have locations zero and scale parameters 1, and suppose

$$(4.8) \qquad \lim_{N \to \infty} P_N\{c_N(h_{Nj} - \mu_{Nj}) \leq u_j \ , \quad j = 1,\dots,p\}$$

$$= G\left(\frac{u_1 + a_1 B_1}{A_1}, \dots, \frac{u_p + a_p B_p}{A_p} \right.$$

where P_N indicates that the probability is computed for the parameter values $\underset{\sim}{\Delta}_N$ or $\underset{\sim}{\theta}_N$ and where h_{Nj} stands for $h_{Nj}(X_{j1},\ldots,X_{jm(N)};\ Y_{j1},\ldots,Y_{jn(N)})$ and h^*_{Nj} for $h^*_{Nj}(Z_{j1},\ldots,Z_{jN})$. Then for any fixed $\underset{\sim}{\Delta}$ and $\underset{\sim}{\theta}$,

$$\lim_{N\to\infty} P_{\underset{\sim}{\Delta}}\left\{C_N(\hat{\underset{\sim}{\Delta}}_N - \underset{\sim}{\Delta}) \leq \underset{\sim}{a}\right\} = G\left(\frac{a_1B_1}{A_1},\ldots,\frac{a_pB_p}{A_p}\right) \tag{4.9}$$

$$\lim_{N\to\infty} P_{\underset{\sim}{\theta}}\left\{C_N(\hat{\underset{\sim}{\theta}}_N - \underset{\sim}{\theta}) \leq \underset{\sim}{a}\right\} = G\left(\frac{a_1B_1}{A_1},\ldots,\frac{a_pB_p}{A_p}\right) \tag{4.10}$$

Proof: By Theorems 3.2 and 3.4

$$\begin{aligned}
&\lim_{N\to\infty} P_{\underset{\sim}{\Delta}}\left\{C_N(\hat{\underset{\sim}{\Delta}}_N - \underset{\sim}{\Delta}) \leq \underset{\sim}{a}\right\} \\
&= \lim_{N\to\infty} P_{\underset{\sim}{0}}\left\{C_N\,\hat{\underset{\sim}{\Delta}}_N \leq \underset{\sim}{a}\right\} \\
&= \lim_{N\to\infty} P_{\underset{\sim}{0}}\left\{h_j(X_{j1},\ldots,X_{jm(N)};\ Y_{j1}-\frac{a_j}{C_N},\ldots,Y_{jn(N)}-\frac{a_j}{C_N}) \leq \mu_{Nj}\ ,\right. \\
&\qquad\qquad\qquad \left. j=1,\ldots,p\right\} \\
&= \lim_{N\to\infty} P_{\underset{\sim}{0}}\left\{h_j(X_{j1},\ldots,X_{jm(N)};\ Y_{j1},\ldots,Y_{jm(N)}) - \mu_{Nj} \leq 0\ ,\right. \\
&\qquad\qquad\qquad \left. j=1,\ldots,p\right\} \\
&= G\left(\frac{a_1B_1}{A_1},\ldots,\frac{a_pB_p}{A_p}\right).
\end{aligned}$$

This completes the proof for (4.5), and that for (4.6) is completely analogous. Thus combining the results of Theorems 4.1, 4.2 and 4.3 we obtain

THEOREM 4.4. (a) Under the assumptions of Theorem 4.1, $N^{1/2}(\hat{\underset{\sim}{\Delta}}_N - \underset{\sim}{\Delta})$ has asymptotically a p-variate normal distribution with mean zero and covariance matrix $\tau = (\tau_{jk})$ given by

$$(4.11) \qquad \tau_{jk} = \frac{d_{jk}}{\lambda(1-\lambda)e_j e_k} ; \qquad j,k = 1,\ldots,p$$

(b) Under the assumptions of Theorem 4.2, $N^{1/2}(\hat{\underset{\sim}{\theta}}_N - \underset{\sim}{\theta})$ has asymptotically a p-variate normal distribution with mean zero and covariance matrix $\Lambda = (\lambda_{jk})$ given by

$$(4.12) \qquad \lambda_{jk} = f_{j,k}/g_j g_k ; \qquad j,k = 1,\ldots,p$$

Special Cases. A. Let J_j be the inverse of the standard normal distribution, and J_j^* be the inverse of the chi-distribution with one degree of freedom. Then the estimates $\hat{\underset{\sim}{\Delta}}_N$ and $\hat{\underset{\sim}{\theta}}_N$ reduce to the normal scores type estimates $\hat{\underset{\sim}{\Delta}}_{N(\Phi)}$ and $\hat{\underset{\sim}{\theta}}_{N(\Phi)}$ respectively. In this case the covariance matrices $\tau = (\tau_{jk})$ and $\Lambda = (\lambda_{jk})$ reduce to $\tau^{\Phi} = (\tau^{\Phi}_{jk})$ and $\Lambda^{\Phi} = (\lambda^{\Phi}_{jk})$ respectively where

$$(4.13) \qquad \tau^{\Phi}_{jk} = \frac{\gamma_{jk}(F,\Phi)}{\lambda(1-\lambda)B_j(F,\Phi)B_k(F,\Phi)} ; \qquad j,k = 1,\ldots,p$$

and

$$(4.14) \qquad \lambda^{\Phi}_{jk} = \lambda(1-\lambda)\tau^{\Phi}_{jk} ; \qquad j,k = 1,\ldots,p$$

where

$$(4.15) \qquad \gamma_{jk}(F,\Phi) = \begin{cases} 1 & \text{if } j = k \\ \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \Phi^{-1}[F_j(x)]\Phi^{-1}[F_k(y)]dF_{jk}(x,y) & \text{if } j \neq k \end{cases}$$

$$(4.16) \qquad B_j(F,\Phi) = \int_{-\infty}^{\infty} \frac{f_j^2(x)dx}{\phi[\Phi^{-1}(F_j(x))]}, \quad j = 1,\ldots,p$$

where $\phi(\cdot)$ is the density of the standard cumulative normal distribution function $\Phi(\cdot)$. B. Let J_j or J_j^* be the inverse of the rectangular distribution over (0,1). Then the estimates $\hat{\underset{\sim}{\Delta}}_N$ and $\hat{\underset{\sim}{\theta}}_N$ reduce to the Wilcoxon type estimates $\hat{\underset{\sim}{\Delta}}_{N(R)}$ and $\hat{\underset{\sim}{\theta}}_{N(R)}$ respectively. In this case the covariance matrices $\tau = (\tau_{jk})$ and $\Lambda = (\lambda_{jk})$ reduce to $\tau^R = (\tau^R_{jk})$ and $\Lambda^R = (\lambda^R_{jk})$ respectively where

$$(4.17) \qquad \tau^R_{jk} = \begin{cases} [12\lambda(1-\lambda)B_j^2(F,R)]^{-1} & \text{if } j = k \\ \left[\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} F_j(x)F_k(y)dF_{j,k}(x,y) - \frac{1}{4}\right][\lambda(1-\lambda)B_j(F,R)B_k(F,R)]^{-1}, & \end{cases}$$

and

$$(4.18) \qquad \lambda^R_{jk} = \lambda(1-\lambda)\tau^R_{jk} \qquad j,k = 1,\ldots,p$$

where

$$(4.19) \qquad B_j(F,R) = \int_{-\infty}^{\infty} f_j^2(x)dx\,; \qquad j = 1,\ldots,p\,.$$

Now denote

$$(4.20) \qquad \hat{\underset{\sim}{\Delta}}_M = (\hat{\Delta}_{M1},\ldots,\hat{\Delta}_{Mp}) \ , \quad \hat{\underset{\sim}{\theta}}_M = (\hat{\theta}_{M1},\ldots,\hat{\theta}_{Mp})$$

where

$$(4.21) \qquad \hat{\Delta}_{Mj} = \text{median } (Y_{j1},\ldots,Y_{jn}) - \text{median } (X_{j1},\ldots,X_{jm}) \ , \quad j = 1,\ldots,p$$

and

$$\hat{\theta}_{Mj} = \text{median } (Z_{j1},\ldots,Z_{jN}) \ , \qquad j = 1,\ldots,p$$

$$(4.22) \qquad \underset{\sim}{\bar{Y}}_N - \underset{\sim}{\bar{X}}_N = (\bar{Y}_1 - \bar{X}_1,\ldots,\bar{Y}_p - \bar{X}_p) \ , \quad \underset{\sim}{\bar{Z}}_N = (\bar{Z}_1,\ldots,\bar{Z}_p)$$

where

$$(4.23) \qquad \bar{Y}_j - \bar{X}_j = \sum_{\beta=1}^{n} Y_{j\beta}/n - \sum_{\alpha=1}^{m} X_{j\alpha}/m \ ; \quad \bar{Z}_j = \sum_{\alpha=1}^{N} Z_{j\alpha}/N \ .$$

In the next section we shall investigate the performance of the estimates $\hat{\Delta}_N$ and $\hat{\theta}_N$ with respect to the estimates (4.20) and (4.22) based on the vectors of medians and the means respectively. To this end we shall need the following well known result.

THEOREM 4.5. (a) Under the assumption that the variance of marginal distribution function F_j exists and is finite for each $j = 1,\ldots,p$,

$$(4.24) \qquad \lim_{N\to\infty} P_{\underset{\sim}{\Delta}}[N^{1/2}(\underset{\sim}{\bar{Y}}_N - \underset{\sim}{\bar{X}}_N) \le \underset{\sim}{u}] = \Phi_{[\underset{\sim}{0},\Sigma/\lambda(1-\lambda)]}(u_1,\ldots,u_p)$$

where $\Phi_{[\underset{\sim}{0},\Sigma]}$ is the cumulative distribution function of the p-variate normal distribution with mean vector $\underset{\sim}{0}$ and covariance

matrix $\sum = ((\sigma_{jk}))$ and $\sigma_{jk} = \text{cov}\,(X_{j\alpha}, X_{k\alpha})$ when $\underset{\sim}{\Delta} = 0$ or $\underset{\sim}{\theta} = 0$.

(4.25) $$\lim_{N \to \infty} P_{\underset{\sim}{\theta}}[N^{1/2}(\overline{\underset{\sim}{Z}}_N - \underset{\sim}{\theta}) \leq \underset{\sim}{u}] = \Phi_{[\underset{\sim}{0}, \sum]}(u_1, \ldots, u_p)$$

(b) Suppose the marginal distribution $F_j(x)$ is absolutely continuous for each $j = 1, \ldots, p$ and its derivative at 0 denoted by $f_j(0)$ exists and is finite. Then

(4.26) $$\lim_{N \to \infty} P_{\underset{\sim}{\Delta}}[N^{1/2}(\hat{\underset{\sim}{\Delta}}_M - \underset{\sim}{\Delta}) \leq \underset{\sim}{u}] = \Phi_{[\underset{\sim}{0}, \sum^*/\lambda(1-\lambda)]}(u_1, \ldots, u_p)$$

where $\sum^* = ((\sigma^{**}_{jk}))$ is given by

(4.27) $$\sigma^{**}_{jk} = \begin{cases} \dfrac{1}{4f_j^2(0)} & \text{if } j = k \\[2ex] \left\{P_{\underset{\sim}{0}}[X_{j\alpha} > 0,\ X_{k\beta} > 0] - \frac{1}{4}\right\} / f_j(0) f_k(0) & \text{if } j \neq k \end{cases}$$

and

(4.28) $$\lim_{N \to \infty} P_{\underset{\sim}{\theta}}[N^{1/2}(\hat{\underset{\sim}{\theta}} - \underset{\sim}{\theta}) \leq \underset{\sim}{u}] = \Phi_{[\underset{\sim}{0}, \sum^*]}(u_1, \ldots, u_p)\ .$$

5. Asymptotic Efficiency

To obtain an idea of the relative performance of one estimator with respect to another, we employ the notion of "the generalized variance" a concept introduced by Wilks (see Wilks (1962), p. 547). The "generalized variance" of a p-variate random vector $(X_1, \ldots, X_p)$ with non-singular covariance matrix

$\underset{\sim}{\Sigma} = (\rho_{jk}\sigma_j\sigma_k)$ is defined to be var X $= \sigma_1^2 \cdots \sigma_p^2 \det \|\rho_{jk}\|$ where "det" denotes the determinant.

Suppose that the two asymptotically unbiased estimates T and T' with $\underset{\sim}{\theta}$ with asymptotically non-singular matrices $\Sigma^T = (\rho_{jk}^T \sigma_j^T \sigma_k^T)$ and $\Sigma^{T'} = (\rho_{jk}^{T'} \sigma_j^{T'} \sigma_k^{T'})$ require N and N' observations to achieve equal asymptotic generalized variances. Then the asymptotic relative efficiency of T with respect to T' is defined as

$$(5.1) \qquad e_{T,T'} = \lim_{N \to \infty} \frac{N'}{N} = \left[\frac{\sigma_1^{T'} \cdots \sigma_p^{T'}}{\sigma_1^{T} \cdots \sigma_p^{T}} \; \frac{\det \|\rho_{jk}^{T'}\|}{\det \|\rho_{jk}^{T}\|} \right]^{p^{-1}}$$

Now from (4.11), (4.12), (4.24), (4.25), (4.26) and (4.28) the generalized variances of $\hat{\underset{\sim}{\Delta}}_N$, $\hat{\underset{\sim}{\theta}}_N$, $\bar{\underset{\sim}{Y}}_N - \bar{\underset{\sim}{X}}_N$, $\bar{\underset{\sim}{Z}}_N$, $\hat{\underset{\sim}{\Delta}}_M$ and $\hat{\underset{\sim}{\theta}}_M$ are given by

$$(5.2) \qquad \text{var}\,(N^{1/2}\,\hat{\underset{\sim}{\Delta}}_N) = \prod_{j=1}^{p} \frac{d_{jj}}{e_j^2 \lambda(1-\lambda)} \cdot \det \|\rho_{jk}^{\Delta}\|$$

where d_{jk} and e_j^2 are given by (4.1) and (4.2) respectively, and

$$(5.3) \qquad \rho_{jk}^{\Delta} = d_{jk}/d_{jj}^{1/2}\, d_{kk}^{1/2} \;; \qquad j,k = 1,\ldots,p\,.$$

$$(5.4) \qquad \text{var}\,(N^{1/2}\,\hat{\underset{\sim}{\theta}}_N) = \left(\prod_{j=1}^{p} f_{jj}/g_j^2 \right) \det \|\rho_{jk}^{\theta}\|$$

where

$$(5.5) \qquad \rho_{jk}^{\theta} = f_{jk}/f_{jj}^{1/2}\, f_{kk}^{1/2} \;; \qquad j,k = 1,\ldots,p\,.$$

$$(5.6) \qquad \operatorname{var}\,[N^{1/2}(\bar{\underset{\sim}{Y}}_N - \bar{\underset{\sim}{X}}_N)] = \prod_{j=1}^{p} \frac{\sigma_j^2}{\lambda(1-\lambda)} \cdot \det \|\rho_{jk}\|$$

$$= [\lambda(1-\lambda)]^{-p} \operatorname{var}\,(N^{1/2}\,\bar{\underset{\sim}{Z}}_N)$$

$$(5.7) \qquad \operatorname{var}\,[N^{1/2}\,\hat{\underset{\sim}{\Delta}}_M] = \prod_{j=1}^{p} \frac{1}{4\lambda(1-\lambda)f_j^2(0)} \cdot \det \|\rho_{jk}^M\|$$

$$= [\lambda(1-\lambda)]^{-p} \operatorname{var}\,(N^{1/2}\,\underset{\sim}{\theta}_M)$$

where

$$(5.8) \qquad \rho_{jk}^M = \begin{cases} 4[P_o(Z_{j\alpha} > 0,\ Z_{k\alpha} > 0) - \frac{1}{4}] & \text{if } j \neq K \\ 1 & \text{if } j = k \end{cases}$$

Hence denoting $e^{(1)}_{T,T^*}$ and $e^{(2)}_{T,T^*}$ as the efficiency of T with respect to T^* for the one-sample and the two-sample problems respectively, we obtain

$$(5.9) \quad e^{(2)}_{(\hat{\underset{\sim}{\Delta}}_N,\bar{\underset{\sim}{Y}}_N-\bar{\underset{\sim}{X}}_N)} = \left[\prod_{j=1}^{p} \sigma_j^2\ e_j^2/d_{jj}\right]^{p^{-1}} \cdot \left[\det \|\rho_{jk}\| / \det \|\rho_{jk}^{\Delta}\|\right]^{p^{-1}}$$

and similar type of expressions for $e^{(1)}_{(\hat{\underset{\sim}{\theta}}_N,\bar{\underset{\sim}{Z}}_N)}$, $e^{(2)}_{(\hat{\underset{\sim}{\Delta}}_N,\hat{\underset{\sim}{\Delta}}_M)}$ and $e^{(1)}_{(\hat{\underset{\sim}{\theta}}_N,\hat{\underset{\sim}{\theta}}_M)}$. Our main interest is to study the relative asymptotic performances of $\hat{\underset{\sim}{\Delta}}_{N(\Phi)}$, $\hat{\underset{\sim}{\Delta}}_{N(R)}$, $\hat{\underset{\sim}{\Delta}}_M$ and $\bar{\underset{\sim}{Y}}_N - \bar{\underset{\sim}{X}}_N$ in the two-sample case and those of $\hat{\underset{\sim}{\theta}}_{N,\Phi}$, $\hat{\underset{\sim}{\theta}}_{N,R}$, $\hat{\underset{\sim}{\theta}}_M$ and $\bar{\underset{\sim}{Z}}_N$ in the corresponding one-sample case. [At this stage we would also like to draw the attention of the reader to a paper of Bickel (1964) who earlier proposed the estimate $\hat{\underset{\sim}{\theta}}_{N(R)}$ of $\underset{\sim}{\theta}$ and investigated its asymptotic performance with respect to the estimates $\hat{\underset{\sim}{\theta}}_M$ and $\bar{\underset{\sim}{Z}}_N$. Since the estimate $\hat{\underset{\sim}{\theta}}_{N(R)}$ is only a particular member of a relatively more general class of our

estimates, some of our results duplicate his results. However this duplication is retained here in order to provide comprehensive comparison and for the ease of reference.]

Thus applying the definition (5.1), we have

$$(5.10)\quad e^{(2)}_{\hat{\underset{\sim}{\Delta}}_{N(\Phi)},\underset{\sim}{\bar{Y}}_N-\underset{\sim}{\bar{X}}_N} = e^{(1)}_{\hat{\underset{\sim}{\theta}}_{N(\Phi)},\underset{\sim}{\bar{Z}}_N} = \left[\prod_{j=1}^{p} \sigma_j^2\, B_j^2(F,\Phi)\right]^{p^{-1}} \cdot \left[\frac{\det \|\rho_{jk}\|}{\det \|\rho^{\Phi}_{jk}\|}\right]^{p^{-1}}$$

where $B_j(F,\Phi)$ is given by (4.16) and

$$(5.11)\quad \rho^{\Phi}_{jk} = \begin{cases} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \Phi^{-1}[F_j(x)]\Phi^{-1}[F_k(y)]dF_{j,k}(x,y) & \text{if } j \neq k \\ 1 \qquad \text{if } j = k \end{cases}$$

$$(5.12)\quad e^{(2)}_{\hat{\underset{\sim}{\Delta}}_{N(R)},\underset{\sim}{\bar{Y}}_N-\underset{\sim}{\bar{X}}_N} = e^{(1)}_{\hat{\underset{\sim}{\theta}}_{N(R)},\underset{\sim}{\bar{Z}}_N}$$

$$= \left[\prod_{j=1}^{p} 12\sigma_j^2(B_j^2(F,R))\right]^{p^{-1}} \cdot \left[\det \|\rho_{jk}\| \,/\, \det \|\rho^{R}_{jk}\|\right]^{p^{-1}}$$

where $B_j(F,R)$ is given by (4.19) and

$$(5.13)\quad \rho^{R}_{jk} = \begin{cases} 12\left[\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} F_j(x)F_k(y)dF_{j,k}(x,y) - \frac{1}{4}\right] & \text{if } j \neq k \\ 1 \qquad \text{if } j = k \end{cases}$$

$$(5.14)\qquad e^{(2)}_{\hat{\Delta}_{N(\Phi)},\hat{\Delta}_{N(R)}} = e^{(1)}_{\hat{\theta}_{N}(\Phi),N(R)}$$

$$= \left[\prod_{j=1}^{p} B_j^2(F,\Phi)/12B_j^2(F,R)\right]^{p^{-1}} \cdot \left[\frac{\det\|\rho_{jk}^{R}\|}{\det\|\rho_{jk}^{\Phi}\|}\right]$$

$$(5.15)\qquad e^{(2)}_{\hat{\Delta}_{N(\Phi)},\hat{\Delta}_{M}} = e^{(1)}_{\theta_{N(\Phi)},\theta_{M}}$$

$$= \left|\prod_{j=1}^{p} \frac{B_j^2(F,\Phi)}{4f_j^2(0)}\right|^{p^{-1}} \cdot \left|\frac{\det\|\rho_{jk}^{M}\|}{\det\|\rho_{jk}^{\Phi}\|}\right|^{p^{-1}}$$

where ρ_{jk}^{M} is given by (5.8).

When p = 1, the efficiencies (5.11), (5.12), (5.13) and (5.14) reduce to the familiar expressions in the one-sample as well as two-sample univariate case [cf. [5], [6]]. We shall study the above efficiencies in different situations.

Case 1. Total Symmetry. Let us assume that the estimators $\hat{\Delta}_{N(\Phi)}$, $\hat{\Delta}_{N(R)}$, $\hat{\Delta}_{M}$, $\bar{Y}_N-\bar{X}_N$, $\hat{\theta}_{N,\Phi}$, $\hat{\theta}_{N,R}$, $\hat{\theta}_M$ and $\bar{X}_N$ have pairwise independent coordinates. In this situation the covariance matrices reduce to diagonal matrices [cf. [9],[11]] and we have

$$(5.16)\qquad e^{(2)}_{\hat{\Delta}_{N(R)},\bar{Y}_M-\bar{X}_N} = e^{(1)}_{\hat{\theta}_{N(R)},\bar{Z}_N} = 12\left[\prod_{j=1}^{p} \sigma_j^2\ B_j^2(F,R)\right]^{p^{-1}}$$

$$(5.17)\qquad e_{\hat{\Delta}_{N(\Phi)},\bar{Y}_N-\bar{X}_N} = e_{\hat{\theta}_{N(\Phi)},\bar{Z}_N} = \left[\prod_{j=1}^{p} \sigma_j^2\ B_j^2(F,\Phi)\right]^{p^{-1}}$$

$$(5.18)\qquad e^{(2)}_{\hat{\underset{\sim}{\Delta}}_{N(\Phi)},\hat{\underset{\sim}{\Delta}}_{N(R)}} = e^{(2)}_{\hat{\underset{\sim}{\theta}}_{N(\Phi)},\hat{\underset{\sim}{\theta}}_{N,R}} = \frac{1}{12}\left[\prod_{j=1}^{p} B_j^2(F,\Phi)B_j^{-2}(F,R)\right]^{p^{-1}}$$

$$(5.19)\qquad e^{(2)}_{\hat{\underset{\sim}{\Delta}}_M,\bar{\underset{\sim}{Y}}_N-\bar{\underset{\sim}{X}}_N} = e^{(1)}_{\hat{\underset{\sim}{\theta}}_M,\bar{\underset{\sim}{Z}}_N} = 4\left[\prod_{j=1}^{p} \sigma_j^2\, f_j^2(0)\right]^{p^{-1}}$$

Hence from the results of Hodges-Lehmann (1961), Chernoff-Savage (1958) and Mikulski (1963), we have

$$(5.20)\qquad \inf_{F\in\mathfrak{F}} e_{\hat{\theta}_{N,R},\bar{X}_N} = 0.864$$

$$(5.21)\qquad \inf_{F\in\mathfrak{F}} e_{\hat{\theta}_{N,\Phi},\bar{X}_N} = 1$$

$$(5.22)\qquad \inf_{F\in\mathfrak{F}} e_{\hat{\theta}_{N,\Phi},\hat{\theta}_{N,R}} = \frac{\pi}{6}$$

$$(5.23)\qquad \inf_{F\in\mathfrak{F}} e_{\hat{\underset{\sim}{\theta}}_M,\bar{\underset{\sim}{Z}}_N} = 0.33$$

where $\mathfrak{F}$ is the set of all totally symmetric absolutely continuous p-variate distributions.

Now suppose that the components of F are totally symmetric as well as identically distributed, then the efficiencies are independent of p, and hence the same as in the case of one-sample univariate situations.

<u>Case 2</u>. <u>Equally Correlated Distributions</u>. Let us now assume that the distribution of $(X_\alpha^{(j)},X_\alpha^{(k)})$ is independent of j and k when $\underset{\sim}{\theta} = \underset{\sim}{0}$. Then, it turns out that

$$(5.24)\quad e_{\hat{\underset{\sim}{\theta}}_{N(R)},\, \underset{\sim}{\overline{Z}}_N} = e_{\hat{\underset{\sim}{\Delta}}_{N,R},\, \underset{\sim}{\overline{Y}}_N - \underset{\sim}{\overline{X}}_N}$$

$$= 12\, \sigma_1^2\, B_1^2(F,R) \left[\frac{1+(p-1)\rho_{12}}{1+(p-1)\rho_{12}^{*}}\right]^{p^{-1}} \left[\frac{1-\rho_{12}}{1-\rho_{12}^{*}}\right]^{1-p^{-1}}$$

$$(5.25)\quad e_{\hat{\underset{\sim}{\theta}}_{N(\Phi)},\, \underset{\sim}{\overline{Z}}_N} = e_{\hat{\underset{\sim}{\Delta}}_{N(\Phi)},\, \underset{\sim}{\overline{Y}}_N - \underset{\sim}{\overline{X}}_N}$$

$$= \sigma_1^2\, B_1^2(F,\Phi) \left[\frac{1+(p-1)\rho_{12}}{1+(p-1)\rho_{12}^{**}}\right]^{p^{-1}} \left[\frac{1-\rho_{12}}{1-\rho_{12}^{*}}\right]^{1-p^{-1}}$$

$$(5.26)\quad e_{\hat{\underset{\sim}{\theta}}_{N(\Phi)},\, \hat{\underset{\sim}{\theta}}_{N,R}} = e_{\hat{\underset{\sim}{\Delta}}_{N(\Phi)},\, \hat{\underset{\sim}{\Delta}}_{N(R)}}$$

$$= \frac{B_1^2(F,\Phi) B_1^{-2}(F,R)}{12} \left[\frac{1+(p-1)\rho_{12}^{*}}{1+(p-1)\rho_{12}^{**}}\right]^{p^{-1}} \left[\frac{1-\rho_{12}^{*}}{1-\rho_{12}^{**}}\right]^{1-p^{-1}}$$

$$(5.27)\quad e_{\hat{\underset{\sim}{\theta}}_{M},\, \underset{\sim}{\overline{Z}}_N} = e_{\hat{\underset{\sim}{\Delta}}_{M},\, \underset{\sim}{\overline{Y}}_N - \underset{\sim}{\overline{X}}_N} = 4\sigma_1^2\, f_1^2(0) \left[\frac{1-\rho_{12}}{1-\rho_{12}^{M}}\right]^{1-p^{-1}} \left[\frac{1+(p-1)\rho_{12}}{1+(p-1)\rho_{12}^{M}}\right]^{1-p^{-1}}$$

<u>Case 3</u>. <u>Normal Case</u>. Let the underlying distribution function F be a non-singular p-variate normal distribution function with mean vector zero, and covariance matrix $\underset{\sim}{\Sigma} = (\sigma_{jk}\sigma_j\sigma_k)$ then, from (5.12), (5.10), (5.14) and (5.15) we obtain

$$(5.28)\quad e_{\hat{\underset{\sim}{\theta}}_{N(R)},\, \underset{\sim}{\overline{Z}}_N} = e_{\hat{\underset{\sim}{\Delta}}_{N(R)},\, \underset{\sim}{\overline{Y}}_N - \underset{\sim}{\overline{X}}_N} = \frac{3}{\pi} \left[\frac{\det \|\rho_{jk}\|}{\det \|\rho_{jk}^{*}\|}\right]^{p^{-1}}$$

where

$$(5.29) \qquad \rho^*_{jk} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \Phi_j\left[\frac{x}{\sigma_j}\right] \Phi_k\left[\frac{y}{\sigma_k}\right] d\Phi_{j,k}(x,y) \ .$$

$$(5.30) \qquad e_{\hat{\underset{\sim}{\Delta}}_{N(\Phi)},\, \underset{\sim}{\bar{Y}}_N - \underset{\sim}{\bar{X}}_N} = e_{\hat{\underset{\sim}{\theta}}_{N,\Phi},\, \bar{Z}_N} = 1$$

$$(5.31) \qquad e_{\hat{\underset{\sim}{\theta}}_{N(\Phi)},\, \hat{\underset{\sim}{\theta}}_{N(R)}} = \frac{\pi}{3}\left[\frac{\det \|\rho^*_{jk}\|}{\det \|\rho_{jk}\|}\right]^{p^{-1}} = e_{\hat{\underset{\sim}{\Delta}}_{N(\Phi)},\, \hat{\theta}_{N(R)}}$$

where ρ^*_{jk} is defined by (5.29),

$$(5.32) \qquad e^{(2)}_{\hat{\underset{\sim}{\Delta}}_{N(R)},\, \hat{\underset{\sim}{\Delta}}_M} = e_{\hat{\underset{\sim}{\theta}}_{N(R)},\, \hat{\underset{\sim}{\theta}}_M} = \frac{3}{2}\left[\frac{\det \|\rho^M_{jk}\|}{\det \|\rho^{\Delta(R)}_{jk}\|}\right]^{p^{-1}}$$

Now Bickel (1964) has proved that for $p > 2$,

$$(5.33) \qquad \inf_{F \in \phi} e_{\hat{\underset{\sim}{\theta}}_{N,R},\, \underset{\sim}{\bar{X}}_N} = 0$$

where ϕ is the set of all non-singular p-variate normal distributions and, since

$$(5.34) \qquad \inf_{F \in \phi} e_{\hat{\underset{\sim}{\theta}}_{N(\Phi)},\, \underset{\sim}{\bar{Z}}_N} = 1 \quad \text{for all } p$$

it follows that

$$(5.35) \qquad \sup_{F \in \phi} e_{\hat{\underset{\sim}{\theta}}_{N(\Phi)},\, \hat{\underset{\sim}{\theta}}_{N(R)}} = \infty \ .$$

Thus we find that when the underlying distribution function is normal, the multivariate normal scores estimator $\hat{\Delta}_{N(\Phi)}$ or $\hat{\theta}_{N(\Phi)}$ (which is as good as the means estimator) can be infinitely better than the multivariate Wilcoxon type estimator $\hat{\underset{\sim}{\Delta}}_{N(R)}$ or $\hat{\underset{\sim}{\theta}}_{N(R)}$ for $p > 2$. This leads to the question of examining the relative performance of these two estimators viz. $\hat{\underset{\sim}{\theta}}_{N(\Phi)}$ (or $\hat{\underset{\sim}{\Delta}}_{N(\Phi)}$) and $\hat{\underset{\sim}{\theta}}_{N(R)}$ or ($\hat{\underset{\sim}{\Delta}}_{N(R)}$) for the case when the underlying distribution function F is bivariate normal $N(0,\sigma_1^2,\sigma_2^2,\rho)$. The efficiency behavior of $\hat{\underset{\sim}{\theta}}_{N(\Phi)}$ and $\hat{\underset{\sim}{\theta}}_{N(R)}$ (or $\hat{\underset{\sim}{\Delta}}_{N(\Phi)}$ and $\hat{\underset{\sim}{\Delta}}_{N(R)}$) is given by the following

THEOREM 5.1. <u>The efficiency of</u> $\hat{\underset{\sim}{\theta}}_{N(R)}$ <u>and</u> $\hat{\underset{\sim}{\theta}}_{N(\Phi)}$ <u>or</u> $\hat{\underset{\sim}{\Delta}}_{N(R)}$ <u>and</u> $\hat{\underset{\sim}{\Delta}}_{N(\Phi)}$ <u>is independent of</u> σ_1 <u>and</u> σ_2 <u>and is given by</u>

$$(5.36)\qquad e_{\hat{\underset{\sim}{\theta}}_{N(R)},\hat{\underset{\sim}{\theta}}_{N(\Phi)}} = \frac{3}{\pi}\left[\frac{1-\rho^2}{1-9(1-\frac{2}{\pi}\cos^{-1}\frac{\rho}{2})^2}\right]^{1/2}$$

$$= \frac{3}{2}\left[-\frac{(1+2\cos\frac{2}{3}(\pi+u))}{u(\pi-u)}\right]^{1/2}$$

<u>where</u> u <u>is determined by</u> $\rho = 2\cos(u+\pi)/3$. <u>The function</u> $e_{\hat{\underset{\sim}{\theta}}_{N(R)},\hat{\underset{\sim}{\theta}}_{N(\Phi)}}$ <u>is monotone increasing for</u> $0 \leq \rho < 1$ <u>and</u> (2) <u>symmetric about</u> $\rho = 0$ <u>and hence unimodal</u>. <u>Finally,</u>

$$(5.37)\qquad \lim_{|\rho|\to 1} e_{\hat{\underset{\sim}{\theta}}_{N(R)},\hat{\underset{\sim}{\theta}}_{N(\Phi)}} = \left(\frac{3}{\pi\sin\frac{\pi}{3}}\right)^{1/2} = 0.91 \quad .$$

The proof is the same as in [2] (see Theorem 5.2 [2]) since when the underlying distribution is normal,

$$e_{\underset{\sim}{\hat{\theta}}_{N(R)},\underset{\sim}{\hat{\theta}}_{N(\Phi)}} = e_{\underset{\sim}{\hat{\theta}}_{N(R)},\underset{\sim}{\bar{X}}_N} .$$

In a similar way the efficiency behavior of $\underset{\sim}{\hat{\theta}}_{N(\Phi)}$ and $\underset{\sim}{\hat{\theta}}_M$ or $\underset{\sim}{\hat{\Delta}}_{N(\Phi)}$ and $\underset{\sim}{\hat{\Delta}}_M$ may be obtained. The details are omitted.

REFERENCES

1. Anderson, T. W., An introduction to multivariate statistical analysis, (1965), John Wiley and Sons, Inc., New York.

2. Bickel, Peter J., On some alternative estimates for shift in the p-variate one-sample problem, Ann. Math. Statist. **35** (1966), pp. 1079-1090.

3. Bloch, Daniel A. and Gastwirth, Joseph L., On asymptotically robust competitors of the one-sample t-test. Technical Report No. 70 (1967), Department of Statistics, The Johns Hopkins University.

4. Gastwirth, Joseph L., On robust procedures, Journ. Amer. Statist. Assoc., **61** (1966), pp. 929-948.

5 Chernoff, Herman and Savage, L. R., Asymptotic normality and efficiency of certain nonparametric tests, Ann. Math. Statist. **29** (1958), pp. 972-994.

6. Hodges, J. L., Jr. and Lehmann, E. L., Estimates of location based on rank tests, Ann. Math. Statist. **34** (1963), pp. 598-611.

7. Johnson, N. L. and Leone, F. C., Statistics and experimental design, (1964), John Wiley and Sons, Inc., New York.

8. Mikulski, Piotr Witold, On the efficiency of optimal non-parametric procedures in the two-sample case, Ann. Math. Statist. **34** (1963), pp. 22-32.

9. Puri, Madan L. and Sen, Pranab K., On a class of multivariate multisample rank order tests, <u>Sankhya</u> **28** (1966), pp. 353-376.

10. Rao, C. R., <u>Linear statistical inference and its applications</u>, (1965), John Wiley and Sons, Inc., New York.

11. Sen, Pranab K. and Puri, Madan L., On the theory of rank order tests for location in the multivariate one-sample problem, <u>Ann. Math. Statist</u>. **38** (1967) (in press).

12. Siddiqui, M. M., Distribution of quantiles in samples from a bivariate population, Journ. Research of the N.B.S. 64 (1960), pp. 145-150.